Generating Music, AI and Energy

MARCIN STRZELECKI
The Krzysztof Penderecki Academy of Music in Kraków
Email: marcin.strzelecki@amuz.krakow.pl

sciendo

Generating Music,
AI and Energy

Musicology Today • Vol. 22 • 2025
DOI: 10.2478/muso-2025-0003

ABSTRACT

This article explores the intersection of artificial intelligence and music, focusing on the conceptual foundations of AI in music generation. It underscores the technological advancements and limitations of AI in capturing the nu-anced elements of traditional music composition. The article addresses the broader implications of utilizing AI, in-cluding the significant energy consumption involved in training and operating these models. It critically examines the sustainability of AI-driven music production against the backdrop of cultural, ecological, and economic consid-erations. By providing insights into both the technical and aesthetical aspects of AI in music, the article offers an overview that aids in understanding the potential and challenges of integrating AI into the creative process, while al-so considering the evolving role of music in human societies.

Keywords: artificial intelligence, music generation, cultural authenticity, energy consumption, autonomous procedures

PEOPLE CREATE, MACHINES GENERATE

For many reasons, it is worth maintaining a distinction between creating and generating. We do not say that Beethoven 'generated' his symphonies or that Bach 'generated' his fugues. Similarly, we could avoid much confusion by refraining from using phrases like 'a computer program created music' or 'can an AI model create music?' If traditionally understood 'creation' is an intentionally controlled action, then 'generation' is an objective and spontaneous process, resembling the workings of nature. In the latter case, we are talking about the phenomenon of generativity, where creators delegate a part of the work to a specific procedure, which autonomously shapes selected aspects of the piece. Although generativity is often associated with some form of technology, it is, in essence, a way of producing artifacts.[1]

GENERATING MUSIC

Generating artifacts, such as literary works, images, or music, is not a new concept. The fascination with generation can be linked to myths and symbols of life. Creating life and beauty leads to rare forms that would not arise by chance. Shaping matter requires intention and energy. There is a low probability that shaking a box of chess pieces will arrange them by type and color, but introducing order requires external energy.[2]

The history of music provides numerous examples of fascination with generativity. These include Guido of Arezzo's *Micrologus* (Chapter 17, eleventh century[3]), Athanasius Kircher's *arca musarithmica* (from *Musurgia universalis,* 1650), and the eighteenth-century musical dice game.[4] Additionally, Goethe's idea of the *Urpflanze* and the later Schenkerian concept of 'deep structure' illustrate the idea of simple rules or procedures leading to complex entities. Similarly, the canonic transformations in the music of the Renaissance and Baroque periods, as well as the serial techniques of the twentieth century, illustrate this fascination.[5]

ELECTRICITY AS A BREAKTHROUGH

The intense development of generativity is influenced by many factors, one of the most significant of which is certainly the widespread availability of electricity. Sound is a type of movement of air molecules, and producing sound requires kinetic energy. Traditionally, this energy was supplied by humans, who used energy from food to make music. This changed in the twentieth century. The first electric devices for reproducing recorded music, electronic instruments, electronic music studios with their equipment, computers, and finally artificial intelligence technology – all became possible thanks to access to electricity. Ultimately, a speaker's membrane moves due to the electric current flowing through an electromagnetic coil.[6]

1 There are many etymological examples that support this statement. Compare, for instance, the Christian *Credo*: *γεννηθέντα οὐ ποιηθέντα* (Greek), and *genitum non factum* (Latin).

2 See the second law of thermodynamics and the concept of entropy.

3 Guidonis Monachi Aretini, *Micrologus ad prestanitores codices MSS. exactus* (Rome,1904), 40–42.

4 See Gerhard Nierhaus, *Algorithmic Composition: Paradigms of Automated Music Generation* (Vienna, 2010), 36.

5 See Marcin Strzelecki, *Niech się stanie muzyka!: O muzycznej generatywności* (Kraków, 2024), 59–60.

6 Alexander Graham Bell obtained a patent for the 'electro-mechano-acoustic' transducer in 1876, see: Mieczysław Kominek, *Zaczęło się od fonografu* (Kraków, 1986), 11.

THE AESTHETIC BREAKTHROUGH

The second factor in generativity's development are the aesthetic and ideological shifts in the early twentieth century. Romantic individualization clashed with bourgeois expectations, leading to the anti-art movement as a response to the era's turbulence. This challenged the Grand Theory of Beauty, which was based on sensory perception and symmetry, instead bestoweing ethical value on what had been considered 'ugly', as a higher order of beauty.[7] Generating music free from traditional beauty – without regularity[8] or clear proportions – is easier, whereas creating complex forms like Mozart's symphonies is challenging.[9] Generativity fascinates by prioritizing procedural actions over outcomes.

In the twentieth century, chance became accepted in the arts on its own terms, with aleatory compositional techniques as a mark of the era. Irregularity, randomness, and generativity are linked, as seen in process-oriented art, where procedure holds meaning and aesthetic value, exemplified – among many others – by John Cage's use of the Yìjīng.

TYPES OF PROCEDURAL OUTCOMES

Based on the type of outcome, I propose to distinguish three types of procedures, which result in:

– any sounds (including their absence),
– sounds exhibiting noticeable regularities,
– very precisely defined sounds that imitate an existing pattern, such as those resembling a Mozart symphony.

This is a simplified classification, but it will work as a starting point. In the first case, the result is unrestricted, beyond the limits of audibility or moderate volume. In this case the procedure, like those of experimental composers of the twentieth century, attracts attention and can be aesthetically appealing, or philosophically motivated. In the second case, sounds approximate the traditional understanding of music by exhibiting regularities, such as scales or metric patterns, and audible transformations. Artificial intelligence can be used, but a simple programming code will suffice.[10]

It is only in the last case that the task is to develop an autonomous procedure that results in a convincing imitation of traditional music. This refers not only to classical European music but also to any music that expresses a certain culture (more on this below). To put it differently, a procedure capable of performing this task should be classified under artificial intelligence, as music of this type, when created by humans, requires some form of intelligence.[11]

GENERATIVE AI AND MUSIC

The development of generative AI models accelerated with the advent of the 'generative pretrained transformer' (GPT) architecture, which introduced the *attention* mechanism.[12] The model generates the most probable continuation of symbols in context by using multi-head attention, allowing it to access previously generated symbols.

New generative models have changed the paradigm of algorithmic music production. Previously, systems focused on separating components such as melody, harmony, and rhythm, similarly as it is done by music

7 Compare the description of 'unpleasant' consonances in: Arnold Schönberg, *Harmonielehre* (Vienna, 1922), 391–392.

8 Such an interpretation of the coherence of sensory stimuli can be found in: Piotr Winkielman, Michał Ziembowicz, and Andrzej Nowak, 'The coherent and fluent mind: How unified consciousness is constructed from cross-modal inputs via integrated processing experiences', *Frontiers in Psychology*, 6/83 (2015), 1664–1078.

9 In the sense of *effective complexity*, cf. Gell-Mann Murray and Lloyd Seth, 'Effective Complexity', in Harald Fritzsch, ed., *World Scientific Series in 20th Century Physics: Volume 40: Murray Gell-Mann. Selected Papers* (World Scientific, 2010).

10 The multimedia examples page for the book *Niech się stanie muzyka!* (op. cit.) contains such examples along with a programming code, e.g., at https://media.amuz.krakow.pl/?a=losowanie, accessed 16 May 2025.

11 Another question is whether we are discussing artificial intelligence or merely simulated intelligence. For further exploration, see more on the so-called 'Turing test' (Alan M. Turing, 'Computing Machinery and Intelligence', *Mind*, LIX/236 (1950), 433–460) as well as recent comments by Penrose (Stuart Hameroff and Roger Penrose, 'Orchestrated reduction of quantum coherence in brain microtubules: A model for consciousness', *Mathematics and Computers in Simulation*, 40/3 (1996), 453–480) and John Searle (John R. Searle, 'Minds, brains, and programs', *Behavioral and Brain Sciences*, 3/3 (September 1980), 417–424).

12 Ashish Vaswani and others, 'Attention Is All You Need', *Advances in neural information processing systems*, 30 (2017), https://doi.org/10.48550/arXiv.1706.03762, accessed 16 May 2025.

theorists.[13] GPT models do not operate like this. Instead, they do what children do: learn through observation. There are no strict rules or algorithmic prescriptions here. The network operates stochastically and heuristically.

ENERGY AND MATERIAL RESOURCES

Training AI models and their generative operation consume enormous amounts of energy,[14] achieved through the parallel operation of many processors. For example, the NVIDIA A100 accelerator, with around 7000 CUDA cores, consumes up to 400 W of power. Training a model like GPT-4 requires about 25,000 such devices for at least three months, consuming tens of thousands of megawatt-hours of energy, enough to power about a thousand households for a year. This consumption increases as many companies and users simultaneously train and use AI models.

Data on global energy consumption by AI is approximate and quickly becomes outdated. In 2023, Alex de Vries estimated that a single Google query costs 0.3 Wh, while a ChatGPT (gpt-4) query costs 3 Wh. When Google search engines generate AI responses, consumption rises to 9 Wh per query, 30 times more than a standard search.[15] AI accounts for a significant portion of energy consumption in big tech, comparable to the consumption of entire countries. Microsoft plans to reactivate the Three Mile Island nuclear power plant to power its server farms and AI model training.[16]

One argument for AI development is its potential to solve global issues, such as the climate or energy crisis, yet this remains hypothetical. Meanwhile, energy consumption and ecological problems are escalating. Reliance on future technological development is a risky gamble.[17] Additionally, IT technology has a substantial material footprint. Producing a small computer requires about 800 kg of raw materials.[18] Production also generates waste, which often ends up in the poorest countries. We must not forget about the water footprint, which could rise to 6.6 billion cubic meters by 2027 used up for cooling AI devices.[19]

An intriguing alternative is *in-materio* computing. For example, appropriately prepared nanomaterials can be used to process acoustic information to distinguish dissonance from consonance.[20] We are increasingly aware that the nature around us is a complex system for processing information. In the case of fungal networks[21] or slime mold colonies,[22] analogies to biological computers are strikingly clear.

HOW ENERGY-INTENSIVE IS AI MUSIC GENERATION?

In terms of music production in computing machines, there are no official estimates of energy consumption provided by main AI companies. It is assumed that energy use would be similar to that needed for image generation.[23] Generating an image requires an amount

13 Dorien Herremans, Ching-Hua Chuan, and Elaine Chew, 'A Functional Taxonomy of Music Generation Systems', *ACM Computing Surveys*, 50/5 (2017).

14 Camilla Hodgson, 'Booming AI demand threatens global electricity supply', *Financial Times*, 17 April 2024, https://www.ft.com/content/b7570359-f809-49ce-8cd5-9166d36a057b, accessed 16 May 2025.

15 Alex de Vries, 'The growing energy footprint of artificial intelligence', *Joule*, 7/10 (2023), 2191–2194.

16 Richard Luscombe, 'Three Mile Island nuclear reactor to restart to power Microsoft AI operations', *The Guardian*, 20 Sept. 2024, https://www.theguardian.com/environment/2024/sep/20/three-mile-island-nuclear-plant-reopen-microsoft, accessed 16 May 2025.

17 On this topic, see: chapter 'The thing about technology' in Jason Hickel, *Less is more: How degrowth will save the world* (London, 2020).

18 Rebeca Grynspan, *Foreword*, in *Digital economy report 2024: shaping an environmentally sustainable and inclusive digital future* (Geneva, 2024), V.

19 Pengfei Li, Jianyi Yang, Mohammad A. Islam, and Shaolei Ren, 'Making AI Less "Thirsty": Uncovering and Addressing the Secret Water Footprint of AI Models', *arXiv*, 26 March 2025, https://doi.org/10.48550/arXiv.2304.03271, accessed 16 May 2025.

20 Dawid Przyczyna, Piotr Zawal, Tomasz Mazur, Marcin Strzelecki, Pier L. Gentili, and Konrad Szaciłowski, 'In-materio neuromimetic devices: dynamics, information processing and pattern recognition', *Japanese Journal of Applied Physics*, 59/5 (2020), 050504.

21 See the special volume in the Springer series *Emergence, Complexity and Computation*: Andrew Adamatzky, ed., *Advances in Unconventional Computing*, vol. 23 (Cham, 2017).

22 Edward Braund and Eduardo Reck Miranda, 'On Building Practical Biocomputers for Real-world Applications: Receptacles for Culturing Slime Mould Memristors and Component Standardisation', *Journal of Bionic Engineering*, 14/1 (2017), 151–162.

23 See estimates on the carbon footprint of image generation in: Alexandra Sasha Luccioni, Yacine Jernite, and Emma Strubell, 'Power Hungry Processing: Watts Driving the Cost of AI of energy comparable to charging a smartphone,[24] not including the energy-intensive process of model training. There is no information online about the energy used by popular services. Thus, obtaining precise data would require conducting measurements on personal computers, which is beyond the capabilities of musicology institutes.

However, some estimates can be obtained based on self-collected data and general officially disclosed information. Currently, the only reliable data we can obtain come from our own experiments. For example, a very simple artificial neural network[25] trained to imitate the distribution of melodic intervals in Palestrina's masses heats the M1 processor up to 35 W for about 5 minutes. This results in approximately 0.42 Wh of electrical energy consumption – enough to power a bike light for half an hour. This may not seem like much, but it provides a reliable starting point for further estimations.

Two published models by Meta labs, AudioGen and MusicGen, are relatively well documented.[26] Generating 30 seconds of very low-quality music based on a text prompt requires about 50% load on an A100 processor (which draws up to 400 W of power) for two minutes, resulting in approximately 3 Wh. This is already six times more than in the previous case. It is important to remember that many people around the world are simultaneously performing similar tasks on their own computers in cloud computing environments. For example, the two models mentioned from Meta labs are downloaded several thousand times a month from the popular repository https://huggingface.co.

Training such models is very costly. The model card for MusicGen on Hugging Face indicates that training takes about a month. The authors state in the documentation that they used 32, 64, and 96 A100 computational units for training the small, medium, and large models, respectively. For the smallest model, this results in roughly 9 MWh.[27]

This is the amount of energy that, on average, two households in the United Kingdom or three households in Poland consume in a year. Adding the medium and large model (two and three times more GPUs) consumes six times more energy – about 55 MWh. The model card contains as many as sixteen published versions (downloadable files), indicating that the models were trained at least about twenty times (including failures). In the case of training AudioGen, 64 and 128 A100 units were used for about a week, resulting in approximately 12 MWh for just one version of the model. This amounts to energy usage that is already very significant – roughly 1 GWh, including both models, sizes, and versions. This can power several hundred households for a year. But this is just the tip of an iceberg.

Similar models are also trained by other institutions and private users. It is important to note that the quality of the musical output from these models is very low. This pertains both to the musical quality, as experienced in online demo versions, and the quality of the sound itself, with its low sampling frequency and many artifacts that degrade the sound.

These models are not designed and implemented for commercial purposes. The programming code is open, allowing individuals to test the amount of energy required to run them. Commercial products (for instance from Suno, Udio, Mustango, Limewire, etc.) are a different matter, and unfortunately, no specific data about them can be found online. However, based on the above basic calculations and supplementing them with information from the internet, we can make some cautious estimates.

First and foremost, it is important to realize that commercially generated results are of significantly better quality. They must be good enough for someone to be willing to pay for them. Meanwhile, a linear increase in quality entails exponential increase in investment (costs of equipment, staff working time, and, of course, energy). Successive models published, for example, by Suno, appear every six months. During this time, many test versions are trained and refined. These models by far exceed the sizes of the aforementioned AudioGen and MusicGen.

Deployment?', *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency*, 2024, 85–99, https://doi.org/10.1145/3630106.3658542, accessed 16 May 2025.

24 Alexandra Sasha Luccioni, Yacine Jernite, and Emma Strubell, 'Power Hungry Processing: Watts Driving the Cost of AI Deployment?', *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency*, 2024, 85–99, https://doi.org/10.1145/3630106.3658542, accessed 16 May 2025.

25 My programming code and a simple result can be found at: https://media.amuz.krakow.pl/?a=palestrina, accessed 16 May 2025.

26 Jade Copet and others, 'Simple and Controllable Music Generation', *Advances in Neural Information Processing Systems*, 36 (2023), 47704–47720, https://doi.org/10.48550/arXiv.2306.05284, accessed 16 May 2025; Felix Kreuk and others, 'AudioGen: Textually Guided Audio Generation', *arXiv*, 5 March 2023, https://doi.org/10.48550/arXiv.2209.15352, accessed 16 May 2025.

27 $32 GPUs \times 0,400 kWh \times 24 h \times 30 days = 9,216 MWh$

However, even at the generation phase, the scale of operation of these models is staggering. In a recent interview, Suno's CEO mentioned twenty-five million users.[28] Assuming that large models consume approximately 5 Wh of power per minute of generated music, and that each user generates only one two-minute piece per week (0.14 per day), this results in 35 MWh of energy per day,[29] or 12.7 GWh annually – enough to power from a few to several thousand households. This is only an estimate for one company and only one piece of music per week per user. Deezer, one of the streaming platforms, estimates that 18% of the music uploaded to its servers is AI-generated (this is a twofold increase year-to-year). This is equivalent to about 20,000 tracks daily, which adds another 150 MWh (under the same assumptions as above) to my figure. The number of published tracks is only a gauge of how much is generated in total.

AI RESULTS IN MUSIC

Despite significant advances in AI technology recently, music generated by such systems still falls short of aesthetic criteria. Naturally, one might quote an old saying here: 'The marvel is not that the bear dances well, but that it dances at all.' Indeed, we were once amazed at the level technology achieved in music generation. Systems like Udio or Suno still leave users astonished upon first encounter. However, after prolonged exposure to such generated music, it becomes evident that machine-made music lacks much when compared to music created by humans. Having familiarised oneself with such generated sounds, one may even identify their specific stylistic features, which can expose them as products of transistors, or at least arouse suspicion.

Additionally, in terms of sound quality, AI-generated works bear the mark of a specialized type of compression. Models do not operate directly on digital sound but on parameters that allow music to be encoded and decoded. This involves so-called neural codecs, such as Encodec (from Facebook labs) and SoundStream (from Google).

28 Peter Yang, *Inside Suno, the AI Music App You Won't Be Able to Stop Listening To – Mikey Shulman (Suno)*, https://creatoreconomy.so/p/inside-suno-the-ai-music-app-you-wont-stop-listening-to-product, accessed 16 May 2025.

29 25,000,000 users×0.14 pieces per day×2 minutes per piece×5 Wh per minute=35 MWh.

A particular challenge for AI models is music that expresses the identity of a social group, as it is the product of a specific culture. Models generate sound creations that are an average of the features of training compositions, with a style that matches user demands. However, the more we narrow our prompt to very specific characteristics, the less the generated music will meet our expectations.

For example, when tasked with imitating a canon or a fugato, these models struggle even with the simplest form of polyphonic imitation between voices in a composition. It is unclear whether this is because the models were not trained on such tasks, or because such pieces were underrepresented in the training dataset, or this is simply a technical problem.

The issue is not just with automatically composing a Baroque fugue or developing themes in a sonata form. Virtually any type of traditionally understood music, which expresses a particular culture's identity, is beyond the reach of AI models. Ask an AI to generate a jod for a raga Marwa, in the teental rhythmic pattern, played on the bansuri with characteristics specific to the Maihar Gharana tradition, and you will see the challenge.

There could be various reasons for this, such as the relatively small representation of specific styles in the training data. Models are trained on very large datasets, often sourced from streaming services. According to recent estimates, platforms such as Spotify and SoundCloud receive over 100,000 new tracks daily, and nearly forty million annually![30] Models learn to generalize training data, approximating the most typical, prototypical features.[31] Meanwhile, much evidence suggests that human musical activity is part of a complex process of cultural inheritance and identity, developed through long-term biological adaptations.[32] We can recognize very subtle musical details, much as we can detect a foreign accent in speech.

30 Glenn Peoples, 'The Ledger: Are There Really 100,000 New Songs Uploaded a Day? Maybe More', *Billboard*, 3 Feb. 2023, https://www.billboard.com/pro/how-much-music-added-spotify-streaming-services-daily/, accessed 16 May 2025.

31 Compare this to the technique of vector quantization (VQ), where the number of data points is reduced to prototype representatives of a given class of objects.

32 Discussion of this issue in: Ruth Anne Rehfeldt, Ian Tyndall, and Jordan Belisle, 'Music as a Cultural Inheritance System: A Contextual-Behavioral Model of Symbolism, Meaning, and the Value of Music', *Behavior and Social Issues*, 30/1 (2021), 749–773.

Recently, social media has seen a trend where people play guessing games about whether a piece of music was composed by a human or a machine. This is a version of the well-known thought experiment called the 'Turing Test'. However, the question whether a given music piece is authentic is poorly framed. Instead, we should ask whether machines can generate this specific type of music today. This question is also incomplete because it lacks specificity concerning the kind of music in question. As I described earlier, different types of sound structures can be generated with varying degrees of success. Generating long-lasting drone sounds without particular harmonic restrictions, or automatically arranging melodies into a canon without regard to resulting harmonies (as seen in Heinz Holliger's *Scardanelli-Zyklus*) is not problematic. However, generating music that a traditional culture identifies with, and that is recognized as authentic by its members, remains an unachievable task.

PUTTING IT ALL TOGETHER

Music production requires energy. In the past, this was primarily kinetic energy needed by instrumentalists and singers to produce sound. Since the twentieth century, electrical energy has moved the speaker's membrane, powered amplifiers and the various devices that generate and process signals into electrical energy. Until recently, energy was also needed for the long process of educating musicians and composers. Today, the experience of many generations of music creators is encapsulated (approximately, as discussed) in the parameters of an artificial neural network. The network models human creative activity through mathematical operations, which are implemented in electrical devices such as silicon-based computational accelerators.

Regardless of the exact amount of energy consumed by AI applications in music, an important question remains concerning the justification for such energy consumption. This issue is certainly intriguing from a scientific perspective. The question of whether the creative process can be formalized and modeled through mathematical operations has sparked the imagination for quite some time. Answers to such questions are significant for philosophical inquiry about the nature of the mind and issues related to the values we hold dear.

Another aspect is viewing the modeling of creativity as a higher stage of modeling cognitive functions, more challenging than 'merely' modeling intelligence. There is a belief that if AI can effectively model creative processes, then artificial general intelligence (AGI) will be within reach. This technology, in turn, is hoped to solve pressing human issues: economic, ecological, food, migration, and others. This explains, among other things, why public attention is so much focused on the advances of AI in fields considered the domain of uniquely human creative activity.

Ultimately, the question of whether AI will be able to 'compose' music in the same way as humans do may never find a definitive answer. Today, machines already generate music that is satisfying for many people and sufficient for various purposes (such as background music for films, advertisements, etc.). Investment in developing neural networks is growing exponentially with the demand. Convincing the last few percent of skeptics might be too costly, both financially and energetically, and the potential benefits may not justify the expenses. As it turns out, it is much easier and cheaper to persuade people to accept the type of music that AI can effectively and relatively inexpensively produce. In other words, AI-generated works might become the baseline or reference point for what people consider music in the future. This is similar to how music reproduced in electronic devices has become the standard compared to music produced directly by instruments. People now treat recorded and played-back sounds as the primary form of music they encounter.

Investing vast energy resources in technology development is driven by the hope of recouping costs through substantial future profits. For those in the classical or artistic music circles, this perspective might not immediately come to mind because the financial stakes in this niche are not as high. However, the entertainment industry is a multi-billion-dollar market.

CONCLUSION

Regardless of future advancements in simulating musical activity through artificial neural networks, it is important to remember that music plays a crucial role in the lives of individuals and communities. As mentioned, it has been with us for a very long time, evolving and adapting to our needs. Edward O. Wilson writes: 'Human beings were made for music. […] Music served early humanity as a means of integrating societies and heightening the emotions of the people […].'[33]

33 Edward O. Wilson, *The meaning of human existence* (New York, 2014), 142.

In comparison, if we were to build a robot that could tirelessly jog, it would not make us – humans – healthier. Recently, when Terry Riley was asked about the future of music, he expressed concern about giving all the fun to the machines.[34] Making music is an important aspect of all human cultures.[35] Conversely, is it possible for cultures to exist without their own music?

REFERENCES

Adamatzky, Andrew, ed., *Advances in Unconventional Computing*, vol. 23 (Cham, 2017).

Aretini, Guidonis Monachi, *Micrologus ad prestanitores codices MSS. exactus* (Rome, 1904).

Braund, Edward, and Reck Miranda, Eduardo, 'On Building Practical Biocomputers for Real-world Applications: Receptacles for Culturing Slime Mould Memristors and Component Standardisation', *Journal of Bionic Engineering*, 14/1 (2017), 151–162.

Brown, Donald E., *Human Universals* (New York, 1991).

Copet, Jade, and others, 'Simple and Controllable Music Generation', *Advances in Neural Information Processing Systems*, 36 (2023), 47704–47720, https://doi.org/10.48550/arXiv.2306.05284, accessed 16 May 2025.

Gell-Mann, Murray, and Seth, Lloyd, 'Effective Complexity', in Harald Fritzsch, ed., *World Scientific Series in 20th Century Physics: Volume 40: Murray Gell-Mann. Selected Papers* (World Scientific, 2010).

Guidonis Monachi Aretini, *Micrologus ad prestanitores codices MSS. exactus* (Rome, 1904).

Hameroff, Stuart, and Penrose, Roger, 'Orchestrated reduction of quantum coherence in brain microtubules: A model for consciousness', *Mathematics and Computers in Simulation*, 40/3 (1996), 453–480.

Herremans, Dorien, Chuan, Ching-Hua, and Chew, Elaine, 'A Functional Taxonomy of Music Generation System, *ACM Computing Surveys*, 50/5 (2017).

Hickel, Jason, *Less is more: How degrowth will save the world* (London, 2020).

Hodgson, Camilla, 'Booming AI demand threatens global electricity supply', *Financial Times*, 17 April 2024, https://www.ft.com/content/b7570359-f809-49ce-8cd5-9166d36a057b, accessed 16 May 2025.

Kominek, Mieczysław, *Zaczęło się od fonografu* (Kraków, 1986).

Kreuk, Felix, and others, 'AudioGen: Textually Guided Audio Generation', *arXiv*, 5 March 2023, https://doi.org/10.48550/arXiv.2209.15352, accessed 16 May 2025.

Li, Pengfei, Yang, Jianyi , Islam, Mohammad A., and Ren, Shaolei, 'Making AI Less "Thirsty": Uncovering and Addressing the Secret Water Footprint of AI Models', *arXiv*, 26 March 2025, https://doi.org/10.48550/arXiv.2304.03271, accessed 16 May 2025.

Luscombe, Richard, 'Three Mile Island nuclear reactor to restart to power Microsoft AI operations', *The Guardian*, 20 Sept. 2024, https://www.theguardian.com/environment/2024/sep/20/three-mile-island-nuclear-plant-reopen-microsoft, accessed 16 May 2025.

Nierhaus, Gerhard, *Algorithmic Composition: Paradigms of Automated Music Generation* (Vienna, 2010).

Peoples, Glenn, 'The Ledger: Are There Really 100,000 New Songs Uploaded a Day? Maybe More', *Billboard*, 3 Feb. 2023, https://www.billboard.com/pro/how-much-music-added-spotify-streaming-services-daily/, accessed 16 May 2025.

Przyczyna, Dawid, Zawal, Piotr, Mazur, Tomasz, Strzelecki, Marcin, Gentili, Pier L., and Szaciłowski, Konrad, 'In-materio neuromimetic devices: dynamics, information processing and pattern recognition', *Japanese Journal of Applied Physics*, 59/5 (2020), 050504.

Rehfeldt, Ruth Anne, Tyndall, Ian, and Belisle, Jordan, 'Music as a Cultural Inheritance System: A Contextual-Behavioral Model of Symbolism, Meaning, and the Value of Music', *Behavior and Social Issues*, 30/1 (2021), 749–773.

Terry Riley, ['on the future of music'], *Ruch Muzyczny*, 19 (2023), 11.

Schönberg, Arnold, *Harmonielehre* (Vienna, 1922).

Searle, John R., 'Minds, brains, and programs', *Behavioral and Brain Sciences*, 3/3 (September 1980), 417–424.

Strzelecki, Marcin, *Niech się stanie muzyka!: O muzycznej generatywności* (Kraków, 2024).

Turing, Alan M., 'Computing Machinery and Intelligence', *Mind*, LIX/236 (1950), 433–460.

Vaswani, Ashish, and others, 'Attention Is All You Need', *Advances in neural information processing systems*, 30 (2017), https://doi.org/10.48550/arXiv.1706.03762, accessed 16 May 2025.

34 Terry Riley, ['on the future of music'] in Agnieszka Szynk-Mika, [questionnaire] 'Pieśń przyszłości zanucą skały?', *Ruch Muzyczny*, 19 (2023), 11.

35 Various forms of music-making are on the list of 200 cultural universals compiled by Donald E. Brown: Donald E. Brown, *Human Universals* (New York, 1991).

Vries, Alex de, 'The growing energy footprint of artificial intelligence', *Joule*, 7/10 (2023), 2191–2194.

Wilson, Edward O., *The meaning of human existence* (New York, 2014).

Winkielman, Piotr, Ziembowicz, Michał, and Nowak, Andrzej, 'The coherent and fluent mind: How unified consciousness is constructed from cross-modal inputs via integrated processing experiences', *Frontiers in Psychology*, 6/83 (2015), 1664–1078.

Marcin Strzelecki. Music theorist and composer, instrumentalist, digital media artist, music critic, graphic artist, calligrapher, electronics enthusiast, and programmer. Affiliated with the Krzysztof Penderecki Academy of Music in Kraków. He lectures on harmony and counterpoint as well as music analysis using computer tools. Apart from composing and performing music, he engages in 'sound art' projects, sound installations, and multimedia projects. He collaborates with the Faculty of Intermedia at the Jan Matejko Academy of Fine Arts in Kraków as well as researchers from the Cracow University of Technology and the AGH University of Science and Technology. He conducts empirically oriented research on the psychological and acoustic conditions of musical creativity, using advanced information processing techniques in his analyses. His research focuses on the perception of musical harmony and timbre, stylometry, and, more recently, the potential applications of artificial intelligence techniques in music creation and analysis. He is the author of a book titled *Let There Be Music! On Musical Generativity*.
ORCID iD: 0000-0002-0103-3053